Entrepreneurial Saving Practices and Reinvestment

Beck, T.H.L.; Pamuk, H.; Uras, R.B.

Publication date:
2014

Document Version
Early version, also known as pre-print

No. 2014-050

ENTREPRENEURIAL SAVING PRACTICES AND REINVESTMENT: THEORY AND EVIDENCE

By

Thorsten Beck, Haki Pamuk,
Burak R. Uras

1 September, 2014

This is also a EBC Discussion Paper
No. 2014-008

ISSN 0924-7815
ISSN 2213-9532

Entrepreneurial Saving Practices and Reinvestment: Theory and Evidence *

Thorsten Beck †
Cass Business School, City University London
CEPR

Haki Pamuk ‡
Tilburg University

Burak R. Uras §
Tilburg University

August 25, 2014

Abstract

We use a novel enterprise survey from Tanzania to gauge the relationship between saving instruments and entrepreneurial reinvestment. While most informal savings practices do not imply a lower likelihood of entrepreneurial reinvestment when compared with formal savings practices, we find a significantly negative effect of *saving within the household* on the likelihood of reinvesting entrepreneurial profits. Our results are robust to an extensive list of robustness checks, including controlling for reverse causation and omitted variable biases. Our work contributes to the recent debate on the implications of different saving instruments in developing countries and expands the entrepreneurial financing constraints literature by focusing on internal rather than external fundings constraints.

Keywords: Entrepreneurial finance; savings; reinvestment; financial inclusion.

JEL Classification: D14; G21; O12; O16.

*We wish to express our gratitude to the Financial Sector Deepening Trust Tanzania (FSDT), for kindly providing us with the main data set used in the study. We would like to thank Daan van Soest, Erwin Bulte, Matreesh Ghatak, Michele Tertilt, Silvio Daidone, Benedikt Goderis, Gonzague Vannorenberghe, Bert Willems, Janneke Pieters, Christian Mpalanzi and FSD Tanzania, and seminar and conference participants at Tilburg Economics Workshop, 2014 Tilburg Development Economics Workshop, and CSAE Conference 2014 for valuable comments and suggestions. Haki Pamuk also would like to thank N.W.O. for financial support (N.W.O. grant number 453-10-001).This research was funded with support from the Department for International Development (DFID) in the framework of the research project 'Coordinated Country Case Studies: Innovation and Growth, Raising Productivity in Developing Countries'. All remaining errors are ours.

†E-mail: TBeck@city.ac.uk

‡E-mail: h.pamuk@uvt.nl

§Corresponding Author. E-mail: r.b.uras@uvt.nl

1 Introduction

In developing countries, intermediation costs and enforcement frictions constrain access to external finance by entrepreneurs - leaving their earning retention as a key element for small business growth. But, what explains entrepreneurial decisions to reinvest their profits into their own businesses? Given the limited access to formal financial services, many entrepreneurs use informal mechanisms of saving and liquidity management to facilitate their earnings retention, part of a larger phenomenon of informal financing arrangements dominating corporate finance in many developing countries (Allen et al. 2005, 2012). In this paper, we utilize a novel entrepreneurial survey to explore whether entrepreneurial saving practices can explain variation in entrepreneurs' reinvestment decisions. Specifically, we gauge whether the decision to save with formal financial institutions, individually (under the mattress), within the household or via other informal arrangements, such as rotating savings and credit associations (ROSCAs), affect the decision to reinvest entrepreneurial earnings. We motivate our empirical work with a simple theoretical model that shows that an entrepreneur's reinvestment decision depends on the entrepreneur's saving practice, in addition to productivity and borrowing capacity of her entrepreneurial firm.

In the absence of easy access to external finance, saving for business purposes should be positively correlated with profit reinvestment decisions. However, the saving mechanism itself might be a critical element in determining the ability to reinvest. On the one hand, for formal savers the opportunity cost of consuming savings instead of reinvesting them is not only the loss of financial reserves but also the foregone interest income. On the other hand, the "within-household savers" might be less likely to reinvest, because they suffer from the redistributive pressure resulting from the saved funds being held inside the household. If the remaining household members are aware of the existence of entrepreneurial savings, it can be hard to prevent the funds from being exploited for the general consumption needs of the household. In addition to these two extreme cases, we can also think of the "individual savers" and the "informal finance network savers"as alternative saving practice types. Comparing "individual savers" with "informal network savers", we note that although the interest income from informal finance networks should have a positive impact on the opportunity cost of consumption and foster investment, the inflexibility to withdraw savings at informal financial institutions

might offset this income effect and reduce the earnings retention.[1]

In order to inform our empirical hypotheses we first present a simple theoretical model to explain the relationship between entrepreneurial investment decisions and saving practices. We show that entrepreneurs are more likely to invest in their businesses if they save in a fashion which allows them easy access to their funds, such as formal savings accounts or personal saving mechanisms.

To test the empirical relationship between savings patterns and entrepreneurial reinvestment decisions, we use a micro- and small enterprise (MSE) survey for over 6,000 entrepreneurs undertaken in 2010 in Tanzania. The sample of entrepreneurs surveyed covers a large variety of enterprises in different locations, of different gender, educational profile and sectors. We document that entrepreneurs' saving *practices* do indeed co-vary with the likelihood of earnings retention at MSEs. The survey design allows us to differentiate between different savings vehicles, including within household saving, saving under the pillow, informal savings clubs, and formal deposit accounts. Our results reveal that the probability of reinvestment is significantly higher for savers and that when compared against formal deposit account holders, entrepreneurs with informal saving practices are significantly less likely to reinvest. However, the type of informal saving instrument matters in generating an adverse consequence for reinvestment. Specifically, we find that when we compare the practice of keeping savings *within the household* against the practice of *having a deposit account at a formal financial institution*, the latter is more likely to be associated with reinvestment than the former. Other informal savers and formal deposit savers do not exhibit significantly different reinvestment rates.

We conduct a series of checks to ensure the robustness of our results to the inclusion of additional control variables, alternative model specifications, and endogeneity biases. First, stable coefficient estimates across an array of specifications shows that it is unlikely that our results are driven by omitted variable bias (Altonji et al. (2005)). Second, to address the potential reverse causation of high reinvestment on saving practices we utilize the distance to the nearest bank and entrepreneur's age as

[1]The rate of return to savings in social saving clubs is typically lower compared to formal financial institutions. For related discussion see Vonderlack and Schreiner (2002). Entrepreneurs saving via informal channels are more likely to have limited access to their savings. For instance, members of ROSCAs cannot access their savings until their turn comes (see Besley et al. (1993) for a theoretical discussion of ROSCAs), unless there is a relevant secondary market (Calomiris and Rajamaran, 1998). Similarly, moneylenders may postpone repaying the savings or it might be hard to reach them.

instruments in recursive bivariate probit regressions. We use these two instruments, because accessibility to a bank and entrepreneur's age can explain whether the savings will be kept in a bank account or shared with the rest of the household, but - as we show - these two variables are not directly associated with reinvestment decisions. The coefficient estimates in the instrumental variable regressions remain stable and significant across all specifications. Finally, we explore the differential effects of saving patterns on reinvestment decision across groups with different intra-household bargaining power. We find that the negative relationship between saving within the household and reinvestment is stronger for entrepreneurs with lower intra-household bargaining power, such as females and non-household heads.

While research on finance and investment has almost exclusively concentrated on the impact of external finance constraints on investment behavior, our paper adds to a very recent liteature arguing that relaxing the internal finance constraints is vital for investment and enterprise growth in developing countries as much as allowing for access to external finance. In this respect, Banerjee and Duflo (2011) suggest that saving constraints in the developing world could imply poverty traps for low income households. Karlan, Ratan and Zinman (2013) state in their survey article that undersaving can have important welfare consequences, such as variable consumption, low resilience to shocks, and foregone profitable investments.

Our paper has multiple novel contributions for this new financial development research agenda, which focuses on *inefficiency of savings and real economic performance*. First, our paper is the first in identifying that a variety of different informal saving practices are being applied among entrepreneurs. Second, we concentrate on re-investment of entrepreneurial earnings as an outcome variable, as incentives that induce earnings retention or reinvestment are expected to be particularly important in alleviating financial constraints. A large fraction of the existing literature on finance and investment has overlooked this particular type of investment behavior in empirical studies. Third, by exploiting the variation in these two key variables, we show that a large spectrum of the informal saving mechanisms do not decrease the likelihood of entrepreneurial re-investment relative to saving formally. Rather, we show a significant negative impact of one particular informal saving method, namely, saving within the household on re-investment likelihood, and highlight this inefficient saving practice as a

channel through which formal financial development policies could induce entrepreneurial investment in developing countries.

Tanzania is a perfect setting to test the relationship between different saving practices and entrepreneurial investment decisions. Tanzania is a low-income country in East Africa, whose private sector is dominated by micro- and small enterprises. While the financial sector was liberalized in the 1990s and there is a large number of formal financial institutions, access to formal financial services is very low, with only 17% of adults having a formal bank account (World Bank (2012)). Tanzania shares many characteristics with other low-income countries in Africa, including a very disperse population and a high degree of informality.

This paper relates to several distinct literatures. First of all, our study relates to past research on finance and entrepreneurial investment, which has shown that entrepreneurs invest more if they expect high private returns from their investment activity (e.g. Demirguc-Kunt and Maksimovic (1998)). Moreover, there are several studies investigating the impact of access to external finance on investment for microenterprises (Karlan and Zinman (2010a); Karlan and Zinman (2010b); Kaboski and Townsend (2011); Ayyagari et al. (2012); Attanasio et al. (2012) and Banerjee et al. (2013)). Particularly important for our paper is the literature that aims to understand the determinants of *re*-investment among entrepreneurial firms in developing countries and emerging markets. Johnson et al. (2002) examine the relative importance of property rights and external finance in Eastern European countries and find property rights to be relatively more important in explaining the earnings reinvestment among entrepreneurial firms. On the other hand, Cull and Xu (2005) show that access to external finance in the form of bank loans is associated with more reinvestment for Chinese entrepreneurs. We add to this literature by focusing on savings patterns as additional factor to explain the variation in reinvestment decisions across micro- and small entrepreneurs.

Second, our paper relates to a growing literature on the relative importance of formal and informal financing mechanisms in developing countries. Allen et al. (2005, 2012) posit that alternative financing channels, such as internal financing and trade credit as well as informal coalitions of firms, investors, and local governments are more important than formal bank credit in fostering the growth of Chinese and Indian enterprises, while Ayyagari et al. (2010) show that it is bank-finance rather than

informal financing that fosters firm growth in China. Similarly, Fisman and Love (2003) show that industries with higher dependence on trade credit financing exhibit higher growth rates in countries with weak financial institutions. Cull et al. (2009) use a large panel dataset of Chinese industrial firms and find that poorly performing state-owned-enterprises are more likely to redistribute credit to firms with limited access to bank loans. We add to this literature by focusing on the savings side of the debate. We disaggregate informal saving practices and show that some informal saving mechanisms function reasonably well - whereas others do not - in inducing entrepreneurs to reinvest. By this we contribute to understanding where and when the benefits might not exceed the large fixed costs associated with increasing the outreach of formal financial products (Beck, Demirguc-Kunt, and Martinez Peria (2007)).

Third, we contribute to the growing literature concerning the implications of access to different saving instruments in developing countries. There is an increasing number of studies exploring the impact of access to formal banking services on the level of savings (Burgess and Panda (2005); Kaboski and Townsend (2005); Dupas and Robinson (2013a)). A recent experimental study by Dupas and Robinson (2013a) shows that entrepreneurs with formal bank accounts save and invest more in their businesses than entrepreneurs who do not save in formal banks. In a companion study (Dupas and Robinson (2013b)), the authors compare the health investment performance of women saving via various informal saving instruments and find that some of them boost investment in health. Similarly Brune et al. (2013) evaluate the effect of commitment to keep savings accounts on several outcomes for Malawian cash crop farmers. We contribute to this literature by comparing the investment likelihood of formal savers with different types of informal savers such as individual savers, savers via other household members, informal savings club members and moneylenders.

Finally, our paper relates to the literature on barriers to saving in developing countries (see Karlan, Ratan and Zinman (2013), for an overview). In addition to geographic, monetary and regulatory barriers, there are significant social constraints on saving behavior, partly related to the position of the entrepreneur within the household. Previous research has linked participation in informal savings clubs, such as ROSCAs, to intra-household bargaining problems (e.g., Besley et al. (1993), Anderson and Baland (2002)). Social constraints can also explain why entrepreneurs save and borrow at

the same time. Critically, the literature has shown that the relative position within the household is important for saving and investment decisions. For instance, de Mel et al. (2008) show that as the decision making power of women in the household increases, returns to capital and investment for women increase as well. Ashraf (2009) in a lab experiment in Philippines documents that subjects are more likely to save the randomly allocated money in their private deposit accounts if their spouse is not aware of the money, while they prefer to consume if the spouse knows about it. Evidence from an experimental study with 142 married couples in Kenya showed that husbands increase private spending if they receive an income shock. But if their wives receive the shock they do not increase their consumption (Robinson (2011)). Likewise Schaner (2013) finds that well matched Kenyan couples are more likely to use joint accounts instead of costly individual ones. Our study supports these findings by showing that members of the household who have potentially less power in decision making are less likely to turn their household savings into investments.

Unlike many other papers in this literature that discuss randomized control trials (RCTs), our paper relies on cross-sectional survey data and thus faces the usual endogeneity biases. We address these concerns by using instrumental variables and by exploring the differential relationship between savings patterns and reinvestment decision across different entrepreneurial groups. Beyond these methodological differences; however, our analysis also allows a broader exploration of reinvestment decisions across different savings patterns. In addition, we realize that such savings patterns are the outcome of repeated interactions and persistent habits and are thus harder if not impossible to control under a randomized control trial.

The rest of the paper is organized as follows. Section 2 presents a theoretical model to show how saving practices can influence entrepreneurial investment decisions. Section 3 discusses the regression set-up and the set of control variables. Section 4 presents the data we use for our analysis. Section 5 discusses our main findings, while section 6 discusses the determinants of saving choice, tests for reverse causality, provides an extensive list of robustness checks and studies sub-sample heterogeneity concerning our key estimation results. Section 7 concludes.

2 A 2-Period Model

We develop a partial equilibrium heterogeneous firms model to study the interactions between entrepreneurial *business saving practices* and *profit reinvestment*. In our model entrepreneurial heterogeneity has three dimensions: productivity, borrowing capacity, and saving practice. In the benchmark model all of the three dimensions are exogenous. We also extend the benchmark in section 2.4, where we endogenize the saving practice as an entrepreneurial decision.

Parallel to the reinvestment and saving practice variables that we use in the empirical specification in section 3 below, we model reinvestment as a direct re-injection of some of the business profits back into the business. In the following, we first present the economic environment, and then the entrepreneur's maximization problem, before deriving the optimal investment behavior. This allows us to obtain several empirically testable hypotheses.

2.1 Environment

There are two time periods, *1* and *2*; a continuum of entrepreneurs indexed by i; and a good - call it cash - that can be invested, saved or consumed. Entrepreneurs have linear preferences over the life-time consumption such that

$$U_i = c_{1,i} + \beta c_{2,i}, \tag{1}$$

where U is the life-time utility and c_1 and c_2 are consumption levels in period-1 and in period-2 respectively. The parameter β is a discount factor. This linear preference specification is not essential for the qualitative findings of the model. It allows us to solve for the investment likelihood of the entrepreneur as we will present in equations (11) and (12) below.

In the beginning of period-1, each entrepreneur is endowed with ω units of investable funds which we assume to be homogeneously distributed among all entrepreneurs in the economy. We interpret this endowment as the net earnings generated by the entrepreneur a priori to period-1. In period-1, entrepreneurs can utilize the ω-endowment to *re*-invest as productive capital ($k_{1,i}$), save

as liquid reserves for business purposes ($s_{1,i}$), or consume ($c_{1,i}$) -yielding the budget constraint for period-1:

$$k_{1,i} + c_{1,i} + s_{1,i} \leq \omega.$$

The uses of productive capital and saving for business purposes are as follows. Each entrepreneur has access to a production technology that converts the investable funds of period-1 into the output of period-2. The output realization of the entrepreneurial technology is conditional on a liquidity injection that needs to be incurred at the beginning of period-2. Specifically, entrepreneur i's technology yields $A_i k_{1,i}$ units of cash in period-2 in return of $k_{1,i}$ units of capital investment in period-1 plus an additional $L(k_{1,i})$ if and only if the entrepreneur is capable of injecting an $L(k_{1,i})$ at the beginning of the period-2 that is greater than $\ell_2 k_{1,i}$. The parameter $A_i > 1$ captures the productivity heterogeneity across entrepreneurs. A high A_i can be associated with better training, education or some sort of intrinsic ability to manage a firm. We assume that A is drawn independently and identically from a distribution at the beginning of the period-1. In this economy, firms must have the capacity to manage liquid reserves in order to be able to undertake productive investment opportunities.[2] The liquidity need of the firm, $\ell_2 k_{1,i}$ poses an inefficient use of capital: Every unit capital saved for liquidity purposes does not get to invested into the productive investment opportunity. The liquidity holdings L can be financed via two sources:

1. The entrepreneur can save cash from period-1 to period-2, which we will call *saving for business purposes* denoted by $s_{1,i}$, at a rate ζ_i with $\zeta_i \leq 1$. In this formulation, ζ_i captures saving practice (in)efficiency of the entrepreneur. We assume that there are two general saving practice types: Formal (ζ_F) and informal (ζ_I) - to be endogenized in section 2.5. We suppose that $\zeta_F = 1$ for those who save formally, whereas ζ_I is drawn from a distribution function with $\zeta_I < 1$. The heterogeneity in informal saving (in)efficiency can be motivated, for instance, by the cross-sectional variation in within-household bargaining power, as we will discuss below.

2. The entrepreneur can borrow, denote it with $b_{2,i}$, up to a θ_i fraction of $\ell_2 k_{1,i}$ in the financial market at a gross interest rate 1, where θ_i is an entrepreneur specific parameter capturing the

[2]We assume that ℓ_2 is a common parameter among all firms in the economy. The qualitative features of the model would remain identical if we assumed heterogeneity and stochasticity in liquidity demand.

ability to raise working capital finance externally. Formally, entrepreneur's borrowing constraint has the following form:

$$\theta_i \ell_2 k_{1,i} \geq b_{2,i}.$$

The borrowing capacity θ is drawn from a distribution function at the beginning of the period-1.

We can state the entrepreneurial total output at the end of the period-2 as follows:

$$\begin{aligned} y_{2,i} &= A_i k_{1,i} + L(k_{1,i}) \ \text{ if } L(k_{1,i}) \geq \ell_2 k_{1,i}, \\ &= L(k_{1,i}) \ \ if \ L(k_{1,i}) < \ell_2 k_{1,i}. \end{aligned} \tag{2}$$

At (2), we make it clear that if and only if the expected liquidity needs in period-2 can be financed, the output available to entrepreneur i will yield the cash-flow generated by the technology, $A_i k_{1,i}$, in addition to the firm's liquidity holdings.[3] Hence, in this economy, firms must have the capacity to manage liquid reserves in order to be able to undertake productive investment opportunities , and savings for business purposes have an indirect impact on reinvestment through liquidity needs.

Entrepreneurs consume and repay using the period-2 output - yielding the period-2 budget constraint

$$c_{2,i} + b_{2,i} \leq y_{2,i}.$$

The timing of events in both periods is specified as the following:

I. **Period-1**

1. Entrepreneurial (3-dimensional) types are realized.

2. Capital investment into the production technology.

3. Saving for business purposes.

[3]This type of a production function specification has been previously utilized in finance and development literature by Aghion et al. (2010): In their dynamic general equilibrium model, the authors introduce a complementarity between the ability to cope with future liquidity needs and current long-term investment and explain the negative correlation between volatility and growth observed in the cross-country data.

4. Period-1 consumption.

I. **Period-2**

1. Borrowing to finance liquidity needs.

2. Liquidity injection: Using borrowed funds and savings from period-1.

3. Cash-flow realization from the production technology.

4. Loan repayment.

5. Period-2 consumption

We would like to note that in this model the exact timing of k investment is not too essential. All we need is that k is invested before the liquidity injection is made. This means allowing parts of the saving for business purposes s to finance k, which can be reached by changing the timing of events (2) and (3) in period-1 timeline, will not alter the qualitative properties of the model that we highlight in section 2.4.

2.2 Optimizing Behavior

The endogenous variables in this model are $c_{1,i}$, $c_{2,i}$, $k_{1,i}$, and $s_{1,i}$. Entrepreneurs maximize life-time preferences delineated at (1) - with respect to the endogenous variables - subject to

$$c_{1,i} + k_{1,i} + s_{1,i} \leq \omega, \tag{3}$$

$$c_{2,i} \leq A_i k_{1,i} + s_{1,i}\zeta_i, \tag{4}$$

where (3) and (4) are the budget constraints for period-1 and period-2 respectively. We would like to note that $\ell_2 k_{1,i}$ enters both sides of the constraint (4); and hence, gets cancelled out. An immediate implication of this model can be summarized with the following.

Lemma 2.1 *If and only if $k_{1,i} > 0$, the entrepreneur forecasts that there will be sufficient capacity to finance future liquidity needs. Therefore, the entrepreneur sets $k_{1,i} = 0$ if his capacity to finance liquidity is sufficiently low.*

This result implies that as long as $k_{1,i} > 0$ we have two additional constraints that need to hold:

$$L(k_{1,i}) \leq s_{1,i}\zeta_i + b_{2,i}, \tag{5}$$

$$\theta_i L(k_{1,i}) \geq b_{2,i}. \tag{6}$$

The qualitative properties of this model are then as follows. Entrepreneurs who choose a $k_{1,i} > 0$, exhaust their borrowing limit θ_i. This is implied by the assumption that saving is inefficient ($\zeta_i < \zeta_F = 1$) in this economy for informal type of saving practices. Therefore,

$$b_{2,i} = \theta_i L(k_{1,i}), \tag{7}$$

as long as $\zeta_i < 1$.

Then using (5) with equality we get:

$$s_{1,i} = \left(\frac{1-\theta_i}{\zeta_i}\right) L(k_{1,i}). \tag{8}$$

Equation (8) implies that the lower ζ the higher is the amount of savings for business purposes - for those entrepreneurs who choose to invest. But, as we show below a low ζ implies a low likelihood of earnings retention and as a result a low likelihood of saving for business purposes.

Using (8) in budget constraints (3) and (4) yields:

$$c_1 = \omega_i - k_{1,i} - \left(\frac{1-\theta_i}{\zeta_i}\right) L(k_{1,i}), \tag{9}$$

$$c_2 = A_i k_{1,i} + (1-\theta_i) L(k_{1,i}). \tag{10}$$

Letting the idiosyncratic rate of return from postponing consumption from period-1 to period-2 be denoted with ρ_i, the optimal consumption plans implied by (1) are described as:

$$\begin{aligned} c_{1,i} > 0,\ c_{2,i} = 0 \text{ if } \rho_i < \frac{1}{\beta}, \\ c_{1,i} = 0,\ c_{2,i} > 0 \text{ if } \rho_i > \frac{1}{\beta}. \end{aligned} \tag{11}$$

Finally, using (9), (10) and the optimal consumption plans from (11) we can show that the entrepreneur chooses to invest ($k_{1,i} > 0$) in period-1 if and only if:

$$\rho_i \equiv \frac{A_i + (1-\theta_i)\ell_2}{1 + \ell_2\left(\frac{1-\theta_i}{\zeta_i}\right)} > \frac{1}{\beta} \tag{12}$$

The left hand side of the inequality (12) is the unit rate of return from undertaking an investment project for an entrepreneur i. The right hand side is the unit cost of postponing consumption from period-1 to period-2. The entrepreneurs with high enough ρ - $\rho_i > 1/\beta$ - invest in their projects and consume the investment returns at the end of the period-2. When ρ_i is lower than $1/\beta$, the entrepreneur does not invest and consumes the endowment ω at the end of the period-1.

2.3 Empirically testable implications of the model

Applying comparative statics at (12) we capture the key empirically testable implication of the model in the following proposition:

Proposition 2.2 *Entrepreneurs with an efficient saving practice (high ζ_i) are more likely to invest.*

Proof Taking the first-partial derivative of ρ with respect to ζ we can see that

$$\frac{\partial \rho}{\partial \zeta} = \left(\frac{1}{\zeta^2}\right)\left\{\frac{\ell_2(1-\theta)(A+(1-\theta)\ell_2)}{\left[1+\ell_2\left(\frac{1-\theta}{\zeta}\right)\right]^2}\right\} > 0,$$

which implies that the rate of return from investing rises with the efficiency of the saving practice of the entrepreneur. A higher efficiency of an entrepreneur's saving practice therefore also raises the likelihood of earnings reinvestment, the key hypothesis of our empirical analysis. □

In order to deepen the empirical validity of our theoretical model, we also provide the following two propositions.

Proposition 2.3 *Entrepreneurs with a high borrowing capacity (high θ_i) are more likely to invest.*

Proof Defining $z \equiv \frac{1-\theta}{\zeta^2}$ and taking the first-partial derivative of ρ with respect to θ:

$$\frac{\partial \rho}{\partial \theta} = \frac{\frac{\ell_2}{z}(A-1)}{[1+\ell_2 \zeta z]^2} > 0,$$

which implies that the rate of return from investing rises with the entrepreneur's borrowing capacity. □

Proposition 2.4 *Productive entrepreneurs (high A_i) are more likely to invest.*

Proof Taking the first-partial derivative of ρ with respect to A

$$\frac{\partial \rho}{\partial A} = \frac{1}{1+\ell_2 \zeta z} > 0,$$

where again $z \equiv \frac{1-\theta}{\zeta^2}$, shows that the rate of return from investment rises with entrepreneurial ability. □

2.4 Endogenizing the Saving Practice

Our theoretical model implies that if an entrepreneur's saving practice is inefficient, then she is induced to save a lot which makes investment, or in other words postponing consumption between period-1

and period-2, harder. Therefore, the saving practice of an entrepreneur is likely to be an endogenous variable, where the decision to save formally might be a costly action.

To formalize this argument, suppose that there are two saving options available for an entrepreneur as spelled out previously - formal and informal. In order to be able to save formally the entrepreneur needs to sacrifice a utility loss worth of ψ_i units of consumption for each unit of fund deposited formally. This basically implies that formal savings impose a non-monetary cost for a class of agents. The utility loss might be due to social costs (e.g. hiding savings from family members at a bank account) or physical costs (e.g. transportation costs) as well as idiosyncratic factors. In addressing the potential reverse causation of investment on entrepreneurial saving practice in section 6, we will utilize entrepreneur's *Age*, Age^2, and *Distance* to the nearest bank as instruments in order to capture the utility loss implied by bank transactions costs.

The efficiency of the formal saving practice is denoted with ζ_F and the efficiency of the informal saving practice is denoted with ζ_I, where $\zeta_F = 1 > \zeta_I$ for all I individuals who save informally. Using equation (12) from the entrepreneurial optimization problem, an entrepreneur i is willing to save formally if and only if

$$\rho_F - \rho_I = (A_i + (1-\theta_i)\ell_2)\left(\frac{1}{1+\ell_2\left(\frac{1-\theta_i}{\zeta_F}\right)} - \frac{1}{1+\ell_2\left(\frac{1-\theta_i}{\zeta_I}\right)}\right) > \psi_i, \tag{13}$$

which would hold if (a) the entrepreneur has a low cost of accessing formal financial institutions and/or (b) a high enough productivity and/or (c) limited access to borrowing.

We utilize the theoretical argument we derived at equation (13), when we study the reverse causation of re-investment likelihood on entrepreneurial saving practice in section 6.

2.5 Impact Heterogeneity

The entrepreneurial (in)efficiency associated with informal saving practices is expected to be a function of accessibility to savings. Such accessibility constraints could be related to the repayment structure

for the case of informal saving networks (e.g. ROSCAs) and to household bargaining power for the case of in-household savings. This implies, for instance, that entrepreneurs with low household bargaining power would have a lower ζ_I. The bargaining power of an individual could vary according to the position of the individual in the household. For instance, due to social norms and pressures female household members, children, and siblings are naturally at a more disadvantageous position than males and household heads in terms of claiming from the common resources of the household. They are less likely to claim money from the common savings pot of the household to finance their liquidity needs and are therefore less likely to reinvest. We will utilize this intuition when studying impact heterogeneity in section 6.

2.6 Empirically Testable Hypotheses

In our regression equations we will control for a vector of variables to test the theoretical results we obtained in propositions 2.1 through 2.3. Specifically, the empirically testable hypotheses resulting from our model are the following:

1. Entrepreneurs who save efficiently (high ζ) are more likely to invest.
2. Entrepreneurs with a high borrowing capacity (high θ) are more likely to invest.
3. Entrepreneurs with better training, higher education and higher income (high A) are more likely to invest.

3 The Empirical Methodology

We test the hypotheses derived from the theoretical model with a dataset collected from Tanzanian MSEs by the Financial Sector Deepening Trust of Tanzania. To test whether saving practices affect the decision to reinvest, we use the binary outcome variable *reinvest*, which equals 1 if the entrepreneur

invests some of the profits back into business, and estimate the following model

$$Reinvest_i = \alpha + \gamma' S_i + \chi' Controls_i + \epsilon_i, \tag{14}$$

where i denotes the entrepreneur, S is a vector of saving practices comprised of dummy variable(s) which take(s) the value of 1 if the entrepreneur has the corresponding saving practice (see below for details) and ϵ is the error term. Since our dependent variable is binary, we estimate probit models for all different specifications of (14), and report marginal effects at mean levels for the coefficient estimates unless we state otherwise. The vector of control variables included in the benchmark model is composed of an array of entrepreneurial and enterprise characteristics that we discuss in the following.

First, in line with our theoretical model, we control for firms' past borrowing history. Specifically, *Borrowed* is a dummy variable which takes the value of 1 if the entrepreneur has ever borrowed to cover business needs, and it is a proxy for the θ_i parameter in the theoretical model. Businesses that have access to external finance are expected to reinvest more frequently even in the absence of regular entrepreneurial savings.

Second, we use income level, education and business training history of entrepreneurs as proxies of entrepreneurial productivity A_i. We conjecture that entrepreneurs with a higher *household income* can save more and as a result reinvest more often. To control for the income effects, we use self-reported monthly personal income levels.[4] Entrepreneurs with a high human capital are expected to be more committed to business growth, and to have higher rates of earnings retention. We therefore use the highest level of *formal education* completed by the respondents, as well as an indicator of entrepreneurial training, as this should matter for expected business performance and reinvestment behavior.

Third, although they are not discussed in our model, we additionally control for *gender* and *marital status* as previous studies showed that both can influence investment decisions (Iversen et al., 2006; Ashraf, 2009; de Mel et al., 2009 and Fafchamps et al., 2013). Specifically, we expect female entrepreneurs to face more claims on their income from spouse and family members. Similarly, married

[4]Each respondent is asked which income range (e.g. TSHS 35 001 - TSHS 40 000 per month) describes their income level best. We use the median of that range (e.g. TSHS 37500.5) as the income level of the respondent.

entrepreneurs might face more claimants on the business profits and might therefore be less likely to re-invest. Finally, we include sectoral dummies to control for sectoral performance that might explain reinvestment heterogeneity, as well as regional dummies to control for geographic heterogeneity in profitability and reinvestment.

We empirically explore the relationship between specific forms of saving and the likelihood of reinvestment. Specifically, our survey allows us to identify two types of saving practices among Tanzanian entrepreneurs which we classify as follows:

1. *Save formal*: This practice includes the entrepreneurs who save their funds at formal financial institutions such as commercial banks, microfinance institutions or saving & credit cooperatives. So entrepreneurs who save only formal and save both formal and informal (please see below for the definitions) means are considered in this group.

2. *Save informal*: We consider entrepreneurs who do not save formally in this group.

This separation corresponds to the control-treatment group set-up of many randomized control trials that assess the impact of using formal savings products on household and entrepreneurial outcomes. In addition, however, our survey allows a finer classification to exploit the considerable heterogeneity in terms of informal saving practices. Therefore we first divide *save informal* into two groups and distinguish individual saving practices and practices involving interaction with other people as follows:

1. *Save informal individually*: A large fraction of entrepreneurs in Tanzania save their funds only in a secret hiding place or piggy bank.[5] We classify this behaviour as "informal individual saving" practice.

2. *Save informal with others*: We classify the practices of saving funds via informal savings clubs, such as ROSCAs, or moneylenders or within household savers under "saving with others". We

[5]Piggy bank is a coin container.

do not include respondents who also save formally in this group. Entrepreneurs who both save informal individually and save informally with others are considered in this group.[6]

To distinguish whether our entrepreneurs save through people living in the household or people who are not member of a household, we decompose the practice of "*Save informal with others*" further into two groups.

1. *Save with household members*: The group comprises of entrepreneurs who give their funds to other household members to keep them safe.

2. *Save with people outside household*: The group contains entrepreneurs who save through ROSCAs or moneylenders. The entrepreneurs who both save informal with household members and save informal with people outside household are considered in this group.[7]

We again conjecture that entrepreneurs in the second group have more control over their savings than entrepreneurs in the first group, especially if the latter have limited intra-household bargaining power. In our regression analysis, we will use a dummy variable for each saving practice above (see Table 1 below for the descriptions) and work with different samples to compare both savers and non-savers but also different groups of savers in their reinvestment behaviour.

4 The Data

The dataset is based on a novel enterprise survey conducted at the MSE-level in Tanzania. The survey data was collected by the Financial Sector Deepening Trust Tanzania in 2010 from a nationwide representative cross-section of 6,083 micro- and small enterprises. The respondents of the questionnaire are entrepreneurs with an active business as of September 2010. Table 1 presents both detailed definitions of the variables and descriptive statistics of the sample.

[6]Our results are robust when we create a separate dummy variable for this group having both saving practices and add them to the regressions.

[7]We do not include the respondents having both practices, saving informal both with people outside household and with household members, to our main regression specifications as only a few respondents (7) do both.

- Table 1 about here -

The descriptive statistics in Panel A of Table 1 shows that the average number of employees among Tanzanian MSEs is 1.5 workers, ranging from one (i.e. self-employed) to 80 employees.[8] However, 97% of entrepreneurs are self-employed. The median initial capital is about 35 USD and average monthly sales are 149 USD. The key question which we exploit to capture entrepreneurs' earnings retention asks whether *the respondent reinvests some of the profits back into business.* As we present in Table 1, 76% of the sample entrepreneurs engage in earnings retention.

The sectoral breakdown in Panel B of Table 1 exhibits substantial variation: 54% and 30% of the businesses operate in the trade and service sectors, respectively, while 15% of enterprises operate in manufacturing.

Panel C of Table 1 presents characteristics of entrepreneurs and enterprises. About 50% of the entrepreneurs in the sample are female, 10% of the entrepreneurs are single. 30% of the sample entrepreneurs received business related training, and about 87% of the entrepreneurs have less than completed secondary education. 75% of the enterprises are located in rural areas. The median monthly personal income of entrepreneurs is 106 USD.[9]

Panel D of Table 1, finally, presents our variables and descriptive statistics on the financing patterns of enterprises in our sample. Only 18% of all sample entrepreneurs ever borrowed for business purposes; 3% of entrepreneurs in the sample borrowed from a bank or MFI, 2% borrowed from a semi-formal financial institution, such as a SACCO or village bank and 6% borrowed from an informal source, such as money lenders, savings club or family and friends.

Saving is a common habit among the entrepreneurs in our sample. We utilize an extensive margin question asking whether the entrepreneur saves for business purposes, and distinguish savers from the rest of the population: 77% of the entrepreneurs in the sample save for business purposes.

[8]The relationship between business owners's saving and re-investment decisions might be weak in large businesses because of managerial layers. We test the robustness of our main result by excluding the businesses larger than 10 from our sample. Estimates reported in Table 3 do not change.

[9]This is computed with the average exchange rate for 2010. If using PPP exchange rates, the corresponding median income would be 288 dollars.

However there is considerable heterogeneity among saving practices of Tanzanian entrepreneurs. Informal individual saving is the most popular practice among Tanzanian entrepreneurs. 75% of the savers save informal-individually whereas around 13% of them save formally. Likewise, 13% of the savers do not save at a formal financial institution and instead save their funds via people outside the household such as members of ROSCAs and moneylenders or give them to household members.

Table 2 presents a correlation matrix concerning the variables of interest for our analysis. The key variables such as "being a saver" and "retaining earnings within the business" exhibit a strong correlation. However, the sign of the relationship seems to be dependent on the saving practice of the respondents. In particular saving via others seems to be negatively correlated with firm reinvestment whereas formal and informal individual savers have higher reinvestment rates. We also note a high correlation among other firm characteristics, such as borrowing and saving activity.

- Table 2 about here -

5 Saving Practices and Reinvestment: Main Results

Table 3 reports the marginal effects for the benchmark regression. We use heteroscedasticity robust standard errors and report the standard deviations associated with coefficient estimates in parentheses.

- Table 3 about here -

The results in the first column show that the probability of reinvestment is higher for both groups of savers compared to non-savers. Specifically, ceteris paribus, the reinvestment probability of an average Tanzanian MSE who saves informally is around six percentage points higher than for an entrepreneur who does not save, while the reinvestment probability of an average Tanzanian MSE who saves formally is around nine percentage points higher. We also find that entrepreneurs with access to formal loans are more likely to reinvest, while formal business training increases the likelihood of reinvestment in business projects. Female and married entrepreneurs are less likely, while richer

entrepreneurs are more likely to invest. Overall, these results are consistent with our theoretical predictions as discussed above and the existing literature.

Our empirical analysis, so far, stresses the significance of *entrepreneurial savings* to foster entrepreneurial reinvestment in business projects and we confirmed that saving related correlations are in line with the findings in the literature. In the next step, we focus on our main research question and we deepen our analysis by studying the implications of saving *practices* on reinvestment. In order to test the predictions from our theoretical model, we rank saving *practices* based on their vulnerability to consumption temptations - as we discussed above - and investigate the implications of the variations in saving methods for the probability to reinvest. Specifically, we rank the "within household savers" as the group for whom the vulnerability to consuming savings is the highest. On the other extreme, we expect the most committed savers to be "formal savers" due to the highest opportunity cost of consumption - resulting from the foregone interest income. Finally, comparing "informal individual savers" with "informal savers with others", we conjecture that while the redistributive pressure problem might be lower for the former, there would be a potential inflexibility to withdrawing savings when needed associated with the latter.

Here we also note that we study our main research question by focusing on specific sub-samples of savers in order to present the results clearer, and keep the consistency between the samples used for main estimations, robustness checks and bivariate probit estimates (see below). To show that our estimates are not biased due to this method, we replicate the analysis by using the entire sample. We present the results in Table A1 in the Appendix, and show that our estimates are robust.[10]

The results in column 2 show that "formal savers" are four percent more likely to retain earnings than the "informal savers". To investigate the effects of individual saving practices on earnings retention we limit our sample to savers and thus drop respondents who do not save. The results in column 3 show that entrepreneurs who save with others are less likely to reinvest than entrepreneurs who save formally. Also, entrepreneurs who save informally but individually are not significantly less likely to reinvest when compared to "formal savers".

[10]The only difference between the results concerns the estimate for save with people outside the household. It is statistically significant at ten percent level due to lower standard error estimates when we use the full sample.

Finally, we focus on the group of respondents who save with others. We independently study the investment likelihood of household savers and respondents who save outside the household compared to the reinvestment probability of formal savers. The regression in column (4) keeps only formal savers and household member savers in our sample, while the regression in column (5) keeps only formal savers and outside household savers in our sample. In both cases, we gauge the difference in reinvestment behaviour relative to formal savers. Therefore, the total numbers of observations in these two regressions are 877 and 774, respectively. Confirming our conjecture, we cannot reject the null hypothesis that "with household member savers" reinvest less frequently compared to "*formal savers*", at the 5% level. Furthermore, we also show that, although the coefficient estimate of *Save with people outside household* variable in the last regression is not significant, the negative coefficient sign is consistent with the argument that the inflexible withdrawal opportunity of "informal savings" might be a barrier to earnings retention.

In summary, our baseline empirical results are consistent with our theoretical model showing that inefficient saving practices lead to lower likelihood of reinvestment. They suggest that informal saving practices are associated with significantly lower likelihood of earnings retention compared to formal saving mechanisms. It is important to note that this finding is mainly driven by the difference in the reinvestment likelihood of within household savers and formal savers, for which the difference is most pronounced and statistically significant.

6 Robustness, Reverse Causality, and Heterogeneity

While controlling for other enterprise and entrepreneurial characteristics reduces the risk that the relationship between savings patterns and the likelihood of reinvestment is a spurious one, we cannot exclude the possibility that our relationship is driven by different sources of endogeneity, including reverse causation and omitted variable bias. As we show in our theoretical model, entrepreneurs who are more willing to reinvest might look for saving practices that support their investment efforts. In this section, we focus on the sample of formal and within-household savers once more since our key result from the empirical analysis of section 5 is that "within household savers" are less likely to re-

invest than "formal savers". Focusing on only one sub-sample also has a methodological advantage as we need fewer exogenous determinants to identify the relationship. We conduct a series of checks to ensure the robustness of our key result concerning the difference in reinvestment likelihood between formal and within household savers (see column (4) in Table 3) to inclusion and exclusion of additional control variables, alternative model specifications and to address endogeneity biases. The section is organized as follows. In subsection, 6.1 we address concerns regarding the non-linear modelling choice. Subsections 6.2 and 6.3 present robustness checks for the omitted variable bias. In subsection 6.4 we utilize an instrumental variable approach and circumvent the reverse causality concerns. Finally, in 6.5 we conduct an analysis to illustrate the heterogeneity of our estimates in sub-samples of data.

6.1 Mis-specification

At first, we address the potential mis-specification of our non-linear probit model. Table 4 provides a set of regression results from an alternative linear model specification - using Ordinary Least Square (OLS) estimation. As linear regression results reported in columns (1)-(5) of Table 4 indicate, our coefficient estimates are consistent and significant, hence stable vis-a-vis our baseline probit regression estimates. Specifically, the OLS estimation shows that saving informally is negatively correlated with the likelihood of earnings retention compared to the formal saving mechanism, and this negative co-variance is significantly driven by the practice of saving with other household members.

- Table 4 about here -

In column 6 of Table 4 we also provide estimation results, where we use a matching model. We do this, because a non-linear correlation between our key explanatory variable - that derives the negative co-variance between saving informal and re-investment (save with household members) - and the control variables may bias our estimates, and we can eliminate this bias by using a matching estimator. By matching each entrepreneur who saves with household members in our study with three (nearest) counterfactual entrepreneurs who save formally (households saving formally), - on the basis of the control variables listed in section 3 and column 4 of Table 3 - we do not assume linear selection

on observables and avoid a bias due to misspecification of the empirical functional form. The results from matching model confirm our previous findings.

6.2 Stability of the Estimates: Removing k-control variables

As a check of omitted variable bias, we provide a stability test for our key regressions with saving with household members in the spirit of Altonji et al. (2005). Specifically, as Bellows and Miguel (2009) propose, we first estimate a parsimonious model and then add control variables gradually. The results in Table 5 show that our marginal effect estimates in columns (1)-(6) are highly stable across different specifications. We measure the stability of marginal effects by calculating the ratio between the value in the regression including controls (numerator) - column (6) - and the difference between this effect and the one derived from a regression without covariates (denominator) - column (1) . As Bellows and Miguel (2009) suggest, this ratio shows how strong the covariance between the unobserved factors explaining entrepreneurial reinvestment decision and savings practices needs to be, relative to the covariance between observable factors and savings practices, to explain away the entire effect we find. The ratio is -4[11], which suggests that to explain the full effect of save within household, the covariance between unobserved factors and savings practices needs to be more than four times as high as the covariance of the included control variables with saving practices.[12] We would like to highlight that in these regressions even adding region fixed effects does not significantly alter our coefficient estimates although adding regional dummies increases pseudo R-square to approximately 4 times, and regional dummies explain most of the variation in the model.

- Table 5 about here -

[11] If we calculate this ratio with OLS instead of probit, as Bellows and Miguel (2009) do, we obtain a ratio of -2.6 – leaving our interpretations unaltered.

[12] In a similar set-up, Altonji et al. (2005) find a ratio of 3.55. The interpretation is that the larger the ratio, the lower the likelihood of unobservables to explain away the entire effect.

6.3 Including Additional Control Variables

As an additional test of omitted variable bias, in Table 6, we test the robustness of our key result with respect to the inclusion of a vector of additional control variables. First, we add specific dummy variables for different sources of external finance at the start-up of the enterprise: formal, semi-formal and informal loans. Our indicator for external finance may not capture the potential implications of access to different sources of finance for reinvestment decisions. Getting loans from a formal financial institution might require a bank account and facilitate formal entrepreneurial savings. However, none of the external financing variables that we include have significant explanatory power for reinvestment likelihood. Second, we control for *entrepreneurial types* by utilizing the answers to the following survey question: "*why did you go to business?*"[13] As evidenced in the previous literature (Bruhn and Zia (2011)), transformational type entrepreneurs are expected to have higher rates of investment profitability and earnings retention rate compared to survival type entrepreneurs. While we do not report the individual dummy variables, some variables enter significant at the 5% level. Third, we add dummy variables to control for the type of the activity the business conducts. The activity of the business (e.g. buying and re-selling; buying, adding value and re-selling, providing a service etc.) may change the definition of re-investment for business owner and timing of the reinvestment. For instance, they may be different for a restaurant owner and a market vendor. To control for this factor, we include answers to the question "what does your business do?" as dummy variables.[14] The estimates for the variables are jointly significant at the 1 percent level. To economize on space we do not report estimates; they are available upon request. Fourth, we include the size of the logarithm of the initial start-up capital, the logarithm of current sales per employee, the logarithm of the duration of business and the logarithm of number of workers since these size gauges are expected to determine the growth potential of a business- and hence the profitability of reinvestment. We also control for rural vs. urban location of the enterprise, as the accessibility to infrastructure might affect expectations

[13]Entrepreneurs selected from a list of statements to indicate why they went into business. Multiple choices were available. The answers include: I was fired / lost/retrenched from a previous job; I couldn't find a job elsewhere; To support me / my family; To try out a business idea; I believe I can make more money working for myself than for someone else; I had nothing else to do/no other means of survival/no better option; parents / relatives were in business; I saw a good opportunity; I have always wanted my own business; I was encouraged by friends and relatives; I needed to supplement my income; Others, please specify.

[14]We include 5 separate dummy variables for the businesses buying and selling goods; buying, adding value and selling goods; making and selling goods; providing service; and other activities including agricultural ones.

and drive variations in reinvestment rates. Including all of these control variables does not affect our key empirical finding.

- Table 6 about here -

Finally, in column (2) we replace the region fixed effects with district fixed effects to ensure that we are capturing geographical variations well enough that could explain the probability of reinvestment. While our sample becomes smaller, our findings remain.[15]

6.4 Instrumental Variable Analysis

In this subsection, we investigate the relationship between saving choices and entrepreneurial characteristics, and then offer a test to alleviate endogeneity concerns. To investigate the determinants of saving choice, we replace the dependent variable *reinvest* with *save within household* in (14) and regress it on our list of control variables as well as on two additional variables denoted by ψ_i in our theoretical model: Age of the entrepreneur and distance to bank. Age increases the bargaining power of the entrepreneur within the household and this implies a U-shaped relationship between age and the choice of within household saving. On the one hand, agents are less likely to be forced to save within household as they get older. On the other hand, when they reach an age giving them enough power to protect their savings within the household, they may be more likely to save with household members. The distance to the nearest bank is expected to increase accessibility of "formal savings services". We estimate two models with two different measures of distance to formal financial institutions. The first one is a subjective distance measure constructed by using the question from the enterprise survey: Is there any bank branch in one hour walking distance to your house? However, there might be a concern regarding the subjective measure, as entrepreneurs who search for formal savings instruments are also those who are more likely to know of the existence of a bank in the close proximity. Therefore, the correlation between the search intensity and some unobserved characteristics may bias our results. For

[15]Note that when we include district fixed effects the total number of observations in the regression decreases to 650 because some districts are excluded from the regression in Probit estimations due to perfect prediction. Our estimates are robust when we estimate the same model with OLS and do not lose any observations.

this reason, we estimate a model with an additional objective distance measure, the logarithm of ward level minimum distance to the closest bank branch, MFI or ATM in 2013 which we constructed using data from the Financial Services Map.[16]

Table 7 reports the marginal effect from probit estimations for the saving practice choice. In columns (1) and (2) of Table 7 we present the results for models including subjective and objective measures, respectively. As we conjecture, the likelihood of saving with household members is higher when entrepreneurs are closer to banks. Moreover, as the age of the entrepreneur increases, he or she is less likely to save with household members. The positive coefficient (0.00038) on the square of age indicates that the age saving with household members practice relationship is non-linear and U-shaped. As the age of the entrepreneur increases, the impact of the age on the saving practice decreases, and getting older increase the probability of saving with household members after the age of 52. The rest of the estimates are also in line with our theory. Entrepreneurs who have access to external finance and entrepreneurs with higher education, better training or high income are more likely to save formally. Finally, female entrepreneurs seem more likely to save in formal institutions - perhaps to escape from redistributive pressures. Also, non-married entrepreneurs are more likely to save formally.

- Table 7 about here -

To address the endogoneity concerns in the relationship between savings practices and entrepreneurial reinvestment decisions, we use an instrumental variable methodology which makes use of the determinants of saving practice choice. Since our dependent and main explanatory variables are binary, we use a system approach, and utilize the age of the entrepreneur and her distance to the nearest bank as instruments in a nonlinear recursive bivariate probit model.[17] Specifically the model

[16]We use data from the Financial Services Map for Tanzania. This data set gives geographic coordinates of bank branches, MFIs and ATMs in 2013 acrossTanzania. We match these data with the existing geographic coordinates of the wards from which entrepreneurial data are collected. Then we calculate the distance of the wards to each financial unit and pick the minimum distance.

[17]We also estimate the same model by using the 2SLS method. We have the same expected signs for the variables of interest but the coefficient estimates are bigger and imprecise as the variance increases. We believe this is because both the dependent and independent variables of interest are binary. Chibus et al. (2012) suggests 2SLS may give very different results and imprecise estimates if the number of observations is lower than 5000 (in our case it is 877).

is formulated as follows:

$$Reinvest_i = \phi + \delta Savehousehold_i + \eta' Controls_i + \sigma_i, \tag{15}$$

$$Savehousehold_i = \lambda + \kappa' Z_i + \pi' Controls_i + u_i. \tag{16}$$

We assume that error terms σ_i and u_i are distributed via bivariate normal distribution. So, $E[\sigma_i] = 0$, $E[u_i] = 0$ and $cov[\sigma_i, u_i] = \mu$. We identify the system by using the vector Z which includes the distance to bank measure and age of the entrepreneur as well as its square and use a similar set of controls as in the main specifications.[18] Table 8 presents the key regressions that we utilize to identify causality between the saving practice and the likelihood of reinvestment.

- Table 8 about here -

There may be exogeneity concerns regarding the relationship between the instruments and reinvestment. For instance, older entrepreneurs may be less likely to reinvest in their businesses. Distance to banks might correlate with business opportunities and induce entrepreneurs to reinvest. In order to address these issues, in columns (1) and (2) of Table 8 we test the exogeneity of our instruments. The standard overidentification test for the 2SLS is not an option with Bivariate Probit estimation. However, it is empirically possible to provide test statistics and argue that our instruments are likely to meet exclusion restrictions. In this respect, at first, we utilize an informal test procedure commonly applied in the empirical literature (e.g. Egger et al. (2011); Booker et al. (2013)), where we introduce our instruments - *Age*, Age^2 and *Distance* - into the benchmark reinvestment regression and study the coefficient estimates. In columns (1) and (2) of Table 8 we show that the coefficient estimates of the instruments are not statistically significant when we include them as additional explanatory variables in our reinvestment regressions together with our main explanatory variable. This implies that the instruments are correlated with saving practices, but they do not co-vary with reinvestment. Second, we show that the estimated cross-correlation coefficient - which we denote with $\hat{\mu}$ in Table 8 - is not statistically significant in either estimation. Therefore, we cannot reject the null hypothesis

[18] We do not use sector dummies in the bivariate probit estimations since our model does not converge. However, not using sector dummies does not change our results since our main results shown in Table 3 are robust when we do not control for them.

that σ_i and u_i are uncorrelated or reinvestment is exogenous to saving practice choice, once we control for other entrepreneurial characteristics. We also test the joint significance of our exogenous variables in the first stage of our bivariate probit model: they are jointly significant at the 1 percent level (Chi-square>20 and p-value<0.001 for both specifications). Finally, we would like to highlight another point concerning the validity of distance measures as instruments. While it is likely that distance to banks - or the density of banks in a region - correlate with local economic environment, which might impact saving practices as well as the local investment opportunities, we control for district fixed effects in our first-stage, which should control for the local economic development level, and this should therefore be a negligible concern for the validity of the distance instrument.

In columns (3) and (4) of Table 8 we present the recursive bivariate-probit estimates by using age in both models, and the two different distance measures as our instruments. Also, Table A2 in the Appendix shows detailed estimation results for the model, including the control variables. The instrumental variable estimations reported in columns (3) and (4) of Table 8 confirm our results. The coefficient estimate of *save with household member* remains negative and significant for both instrument sets. Different measures of distance produce similar results thereby minimizing the concerns regarding the validity of the distance-to-bank proxies. We also note that the estimates for the exogenous variables have the expected signs. The probability to save in the household decreases as the proximity to bank decreases and entrepreneur gets older.

6.5 Heterogeneity

As we discussed in section 2.6, we expect heterogeneous reinvestment responses with respect to the within-household saving practice. Therefore, in order to deepen our analysis and strengthen our identification, we present a set of impact heterogeneity results in Table 9. Specifically, we compare the reinvestment behaviour of entrepreneurs who save with household members with the reinvestment behaviour of entrepreneurs who use formal savings mechanisms across the following two sample splits. First, we split the sample into female and male entrepreneurs. Theory and empirical evidence suggests that social constraints on accessibility of saved funds is higher for women compared to men. Second,

we split the sample into entrepreneurs that are household heads and entrepreneurs that are spouses, children or siblings. We expect the social constraints to be less strong for household heads.

The results in Table 9 confirm the differential relationships between household savings and reinvestment decisions. The results reveal that the marginal effects of *Save with household members* on reinvestment are larger and more significant for female and non-head family members. While the negative relationship between saving within the household and reinvestment decisions are significant at least at the 10% level for all groups, the economic significance is large for female, non-household heads. Supporting our theoretical predictions, this result implies that entrepreneurs who are in disadvantageous positions in their households are more negatively affected from inefficient saving practices.

- Table 9 about here -

7 Conclusion

Past research has identified several factors that are important for entrepreneurial investment in developing countries. In this study, we explored how different entrepreneurial saving practices - i.e. saving via formal financial institutions, individually (under the mattress), within the household or within informal arrangements, such as ROSCAs - are related with the likelihood of reinvestment. To this end, we used a novel survey data set collected from MSEs in Tanzania and distinguished multiple saving practices of entrepreneurs as well their earnings retention behaviour. We motivated our empirical research with a simple theoretical model that shows how different saving practices can influence investment decisions. We have three key empirical results. First, we show that saving and the probability of reinvestment are significantly correlated. Second, we provide evidence that entrepreneurs who save by giving funds to other household members are less likely to reinvest than formal savers. Third, we document that the difference in the likelihood of reinvestment across saving practices is significantly higher for those entrepreneurs who potentially have low bargaining power in the household. We address endogeneity concerns and also provide an extensive list of robustness checks that confirm our results.

Our findings suggest that the entrepreneurs who need to protect their savings from consumption commitments of other household members may benefit most from the introduction of formal saving instruments in low income areas. Therefore, from a development policy perspective, targeting entrepreneurs who have low decision power in the household and facilitating their access to formal saving instruments could be thought as a priority. Our results have important implications for the interactions between enterprise performance and financial access as well. Enterprises that exploit reinvestment opportunities are expected to be more likely to sustain higher productivity levels and survive more often. Access to efficient saving mechanisms in this respect could be key to facilitate enterprise performance in financially developing societies.

Our research raises also some new issues regarding the implications of savings practices of entrepreneurs. First, why do savers inside households not open a bank account to save? Although we implicitly show proximity to banks as an important factor to save in a formal account, identification of all factors is not in the scope of this study. Second, what is the exact role of pressure inside the household that does not allow earnings retention? These important questions we leave to future work.

References

Aghion, P., Angeletos, G.M., Banerjee, A., & Manova, K. (2010). Volatility and growth: Credit constraints and the composition of investment. *Journal of Monetary Economics*, 57 (3), 246-265.

Allen, F., Chakrabarti, R., De, S., Qian, J., Qian, M. (2012). Financing Firms in India. *Journal of Financial Intermediation*, 21, 409-445.

Allen, F., Qian, J., Qian, M. (2005). Law, Finance and Economic Growth in China. *Journal of Financial Economics*, 77(2005), 57-116.

Altonji, J., Elder, T., and Taber, C. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*.

Ashraf, N. (2009). Spousal control and intra-household decision making: An experimental study in the Philippines. *The American Economic Review*, 1245-1277.

Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E., & Harmgart, H. (2011). Group lending or individual lending? Evidence from a randomised field experiment in Mongolia. *EBRD mimeo*.

Ayyagari, M., Demirguc-Kunt, A., Maksimovic, V. (2012). Firm innovation in emerging markets: the role of finance, governance, and competition. *Journal of Financial and Quantitative Analysis*, 46(06), 1545-1580.

Ayyagari, M., Demirguc -Kunt, A., Maksimovic, V. (2010). Formal versus Informal Finance: Evidence from China. *The Review of Financial Studies*, 23(8), 3048-3097.

Banerjee, A., and Duflo, E. (2011). Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. *Public Affairs*.

Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2013). The miracle of microfinance? Evidence from a randomized evaluation. *MIT mimeo*.

Beck, T., Demirguc-Kunt, A., Martinez Peria, M. S. (2007). Reaching out: Access to and use of banking services across countries. *Journal of Financial Economics*, 85(1), 234-266.

Besley, T., Coate, S., & Loury, G. (1993). The economics of rotating savings and credit associations. *The American Economic Review*, 792-810.

Booker, K., Sass, T. R., Gill, B., & Zimmer, R. (2011). The effects of charter high schools on educational attainment. *Journal of Labor Economics*, 29(2), 377-415.

Bruhn, M., & Zia, B. (2011). Stimulating managerial capital in emerging markets: the impact of business and financial literacy for young entrepreneurs. *World Bank mimeo*.

Brune, L., Giné, X., Goldberg, J., & Yang, D. (2011). Commitments to save: A field experiment in rural Malawi. *World Bank Policy Research Working Paper Series*.

Calomiris, C. W., & Rajaraman, I. (1998). The role of ROSCAs: lumpy durables or event insurance? *Journal of Development Economics*, 56(1), 207-216.

Cull, R., Xu, L. C., Zhu, T. (2009). Formal finance and trade credit during China's transition. *Journal of Financial Intermediation*, 18(2), 173-192.

Cull, R., Xu, L. C. (2005). Institutions, ownership, and finance: the determinants of profit reinvestment among Chinese firms. *Journal of Financial Economics*, 77(1), 117-146.

Chiburis, R. C., Das, J., & Lokshin, M. (2012). A practical comparison of the bivariate probit and linear IV estimators. *Economics Letters*.

De Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: evidence from a field experiment. *The Quarterly Journal of Economics*, 123(4), 1329-1372.

De Mel, S., McKenzie, D., & Woodruff, C. (2009). Are women more credit constrained? Experimental evidence on gender and microenterprise returns. *American Economic Journal: Applied Economics*, 1-32.

Demirguc-Kunt, A., & Maksimovic, V. (1998). Law, finance, and firm growth. *The Journal of Finance*, 53(6), 2107-2137.

Dupas, P, & Robinson, J., (2013a). Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics* 5 (1): 163-192. doi:10.1257/app.5.1.163.

Dupas, P, & Robinson, J., (2013b). Why Don't the Poor Save More? Evidence from Health Savings Experiments. *The American Economic Review* 103 (4): 1138-1171.

Egger, P., Larch, M., Staub, K. E., & Winkelmann, R. (2011). The trade effects of endogenous preferential trade agreements. *American Economic Journal: Economic Policy*, 3(3), 113-143.

Fafchamps, M., McKenzie, D., Quinn, S. R., & Woodruff, C. (2011). When is capital enough to get female microenterprises growing? Evidence from a randomized experiment in Ghana (No. w17207). *National Bureau of Economic Research.*

Fisman, R., Love, I. (2003). Trade credit, financial intermediary development, and industry growth. *The Journal of Finance*, 58(1), 353-374.

Iversen, V., Jackson, C., Kebede, B., Munro, A., & Verschoor, A. (2011). Do spouses realise cooperative gains? Experimental evidence from rural Uganda. *World Development*, 39(4), 569-578.

Johnson, S., McMillan, J., Woodruff, C., (2002). Property Rights and Finance. *The American Economic Review*, 92 (5), 1335-1356(22).

Kaboski, J. P., & Townsend, R. M. (2005). Policies and Impact: An Analysis of Village Level Microfinance Institutions. *Journal of the European Economic Association*, 3(1), 1-50.

Kaboski, J. P., & Townsend, R. M. (2011). A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative. *Econometrica*, 79(5), 1357-1406.

Karlan, D., Ratan, A., & Zinman, J. (2013). Savings by and for the Poor: A Research Review

and Agenda (No. 1027).

Karlan, D. & Zinman, J., (2010a.) Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts. *Review of Financial Studies*, 23 (1): 433-64.

Karlan, D. & Zinman, J., (2010b). Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to Estimate the Impacts in Manila. *Unpublished.*

Pande, R., & Burgess, R. (2005). Do rural banks matter? Evidence from the Indian social banking experience. *American Economic Review*, 95(3), 780-795.

Robinson, J. (2012). Limited insurance within the household: evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics*,4(4), 140-164.

Schaner, S. G. (2012). Do Opposites Detract? Intrahousehold Preference Heterogeneity and Inefficient Strategic Savings. *Unpublished manuscript, Dartmouth Coll., Hanover, NH.*

Vonderlack, R. M., & Schreiner, M. (2002). Women, microfinance, and savings: Lessons and proposals. *Development in Practice*, 12(5), 602-612.

Table 1: Descriptive statistics for the main variables

Panel A: Firm characteristics	Description	Obs	Mean	S.D.	Min	Max
Reinvestment	Equals to 1 if respondent re-invest some of the profit back to business, 0 otherwise	6083	0.76	0.43	0	1
Employee	Number of employees business has [including owner]	6083	1.47	1.61	1	80
Initial capital	Logarithm of initial capital of the business, in Tanzanian Shillings	6083	10.62	2.21	0	25.33
Panel B: Sectoral breakdown of firms	Number of companies	%				
Trade	3291	54.1				
Service	1841	30.3				
Manufacturing	931	15.3				
Other	20	0.3				
Panel C: Entrepreneur	Description	Obs	Mean	S.D.	Min	Max
Education	Education level of the respondent, [0 none-6 university]	6077	2.00	0.89	0	6
Female	Equals to 1 if respondent is female, 0 otherwise	6083	0.50	0.50	0	1
Single	Equals to 1 if respondent is single, 0 otherwise	6083	0.10	0.29	0	1
No training	Equals to 1 if respondent has no business related training, 0 otherwise	6083	0.70	0.46	0	1
Rural	Equals to 1 if respondent lives in a rural area, 0 otherwise	6083	0.75	0.44	0	1
Income	Logarithm of personal income level of the respondent in Tanzanian Shillings	5868	11.94	1.15	9.90	15.20
Bank branch within one hour walking distance	Equals to 1 if there is a bank within a one hour walk from the home of the respondent, 0 otherwise	6083	0.30	0.46	0	1
Min. distance to ATM, bank branch, or MFI	Minimum distance of the ward entrepreneur lives to the nearest ATM, bank branch or MFI, in logarithms, at ward level]	583	2.04	1.78	-4.35	6.12
Age	Age of the respondent	6083	36.84	10.6	16	91
Panel D: Finance variables	Description	Obs	Mean	S.D.	Min	Max
Save	Equals to 1 if respondent saves for business purposes, 0 otherwise	6083	0.77	0.42	0	1
Save formal	Equals to 1 if respondent saves in a bank account, MFI or SACCO, 0 otherwise	6083	0.10	0.30	0	1
Save informal	Equals to 1 if respondent saves but not in a bank account, MFI or SACCO and, 0 otherwise	6083	0.67	0.47	0	1
Save informal individually	Equals to 1 if respondent saves in a secret hiding place or piggy bank and does not save via other means, 0 otherwise	6083	0.57	0.49	0	1
Save informal with others	Equals to 1 if savehousehold or saveouthousehold equals to 1 and does not save formally, 0 otherwise	6083	0.10	0.30	0	1
Save with household members	Equals to 1 if respondent save via by giving it to a household member to keep it safe and does not save formally, 0 otherwise	6083	0.06	0.24	0	1
Save with people outside household	Equals to 1 if respondent save via by giving it to a non household member or merry go-round and does not save formally, 0 otherwise	6083	0.05	0.20	0	1
Borrowed	Equals to 1 if respondent has ever taken a loan/ borrowed money for business purpose, 0 otherwise	6083	0.18	0.38	0	1
Formal loan	Equals to 1 if respondent took a to set up or take over the business from a bank or MFI, 0 otherwise	6083	0.03	0.16	0	1
Semi formal loan	Equals to 1 if respondent took a credit from an employer , SACCO, Village Bank, local government schemes or donor/NGO to set up or take over the business , 0 otherwise	6083	0.02	0.13	0	1
Informal loan	Equals to 1 if respondent took a to set up or take over the business from family, friends, savings club, money lender or supplier, 0 otherwise	6083	0.06	0.24	0	1
		6083	0.06	0.24	0	1

Notes: The table shows the varibles definitions and descriptive statistics for the selected variables used in the analysis. Obs., Mean. S.D., Min. and Max. shows the number observation for, sample averagem standard deviation, minimum, and maximum of the corresponding variable.

Table 2: Correlation Matrix: Pairwise correlations among selected variables

	Reinvestment	Save	Save formal	Save informal individually	Save informal with others	Save with household members	Save with people outside household	Borrowed	Education	Female	Single	No training	Income
Reinvestment	1												
Save	0.09*	1											
Save formal	0.07*	0.18*	1										
Save informal individually	0.05*	0.62*	-0.38*	1									
Save informal with others	-0.03*	0.18*	-0.11*	-0.39*	1								
Save with household members	-0.04*	0.14*	-0.08*	-0.29*	0.74*	1							
Save with people outside household	0.00	0.12*	-0.07*	-0.25*	0.64*	-0.03*	1						
Borrowed	0.06*	0.12*	0.30*	-0.09*	0.03*	-0.01	0.05*	1					
Education	0.07*	0.13*	0.24*	-0.04*	0.02	0.02	0.01	0.17*	1				
Female	-0.06*	0.00	-0.05*	0.03*	0.00	-0.10*	0.11*	0.03*	-0.10*	1			
Single	0.02	0.00	0.00	0.01	-0.01	0.00	-0.01	-0.04*	0.07*	0.01	1		
No training	-0.02	-0.06*	-0.08*	0.02	-0.04*	-0.04*	-0.01	-0.10*	-0.12*	0.00	0.00	1	
Income	0.13*	0.12*	0.19*	-0.02	0.01	0.02	-0.02	0.12*	0.19*	-0.24*	-0.02	0.00	1

Notes: The table shows the pairwise correlation coefficient among selected variables shown in first row and first column of the Table. The detailed variable definitions are given in Table 1.
* indicates the estimate is statistically significant at least 5 percent level.

Table 3: Regressions for reinvestment and saving/saving practices relationship

	(1)	(2)	(3)	(4)	(5)
Save formal	0.09***				
	(0.02)				
Save informal	0.06***	-0.04*			
	(0.01)	(0.02)			
Save informal individually			-0.03		
			(0.02)		
Save informal with others			-0.09***		
			(0.03)		
Save with household members				-0.12***	
				(0.04)	
Save with people outside household					-0.04
					(0.03)
Borrowed	0.04**	0.05***	0.05***	0.05	0.04
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Education	0.01	0.01	0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female	-0.03**	-0.03**	-0.03***	-0.07**	-0.04
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)
Single	0.04**	0.03	0.03	0.06	0.03
	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)
No training	-0.03***	-0.04***	-0.04***	-0.07**	-0.06**
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)
Income	0.03***	0.02**	0.02**	0.03**	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	5,803	4,499	4,499	877	774
Sample	All	Savers	Savers	Formal and household savers	Formal and others savers
Base category	No saving	Save formal	Save formal	Save formal	Save formal

Notes: This table shows our baseline estimation results for the relationship between saving practices, control variables and reinvestment likelihood. The detailed variable definitions are given in Table 1. Reinvestment is the dependent variable in the estimations. We estimate Probit models for all specifications in columns 1 to 5 and report marginal effects estimates at mean values for all variables and robust standard errors in parentheses. To control for unobserved regional and sector level fixed effects, we add sector and and region dummies to all estimations. The details of the estimations in the columns 1 to 5 are as follows: (1) In column 1, we compare the reinvestment likelihood of formal and informal savers with non-savers. We use our entire sample for this estimation. The estimate for Save formal (informal) shows the difference between the impact of Save formal (informal) and not saving (base category) on reinvestment likelihood. (2) In column 2, we compare the reinvestment performance of informal savers with formal saver by restricting our sample to only savers and dropping the respondents who do not save. The estimate for Save informal shows the difference between the impact of Save informal and Save formal (base category) on reinvestment likelihood. (3) In column 3, we disentangle informal saving practices to saving informal individually and saving informal with others by adding separate dummies for each group. The estimate for Save informal individually (Save informal with others) shows the difference between the impact of Save informal individually (Save informal with others) and Save formal (base category) on reinvestment likelihood. (4) In column 4, we compare reinvestment likelihood of household savers with formal savers by keeping only formal savers and household member savers in our sample. The estimate for Save informal with household members shows the difference between the impact of Save informal with household members and Save formal (base category) on reinvestment likelihood. (5) In column 5, we compare reinvestment likelihood of outside household savers with formal savers by keeping only formal savers and outside household savers in our sample. The estimate for Save informal with people outside household shows the difference between the impact of Save informal with people outside household and Save formal (base category) on reinvestment likelihood. $p<0.1$. ** $p<0.05$. *** $p<0.01$

Table 4: Estimates for reinvestment and saving/saving practices relationship by using OLS and matching estimators

	(1)	(2)	(3)	(4)	(5)	(6)
Save formal	0.10***					
	(0.02)					
Save informal	0.06***	-0.03*				
	(0.01)	(0.02)				
Save informal individually			-0.02			
			(0.02)			
Save informal with others			-0.08***			
			(0.02)			
Save with household members				-0.11***		-0.12***
				(0.03)		(0.00)
Save with people outside household					-0.04	
					(0.04)	
Borrowed	0.04**	0.05***	0.05***	0.04	0.04	
	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	
Education	0.01	0.01	0.01	-0.01	-0.01	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
Female	-0.03**	-0.03**	-0.03***	-0.06**	-0.03	
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	
Single	0.04**	0.03	0.03	0.06	0.04	
	(0.02)	(0.02)	(0.02)	(0.04)	(0.05)	
No training	-0.03***	-0.04***	-0.04***	-0.07**	-0.06**	
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	
Income	0.03***	0.01**	0.02**	0.03**	0.05***	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
Constant	0.53***	0.54***	0.53***	0.71***	0.31	
	(0.07)	(0.16)	(0.16)	(0.18)	(0.22)	
Observations	5,803	4,499	4,499	877	774	877
R-squared	0.14	0.12	0.13	0.22	0.17	
Method	OLS	OLS	OLS	OLS	OLS	NN Matching
Sample	All	Savers	Savers	Formal and household savers	Formal and others savers	Formal and household savers
Base category	No saving	Save formal	Save formal	Save formal	Save formal	Save formal

Notes: This table shows our baseline estimation results for the relationship between saving practices, control variables and reinvestment likelihood. The detailed variable definitions are given in Table 1. Reinvestment is the dependent variable in the estimations. We estimate OLS models for all specifications in columns 1 to 5 and use propensity score matching method for the estimation shown in column 6. Robust standard errors in parentheses. To control for unobserved regional and sector level fixed effects, we add sector and and region dummies to all estimations. The details of the estimations in the columns 1 to 6 are as follows: (1) In columns 1 to 5, we replicate the estimations shown in Table 3 by using OLS method instead of Probit. we compare the reinvestment likelihood of formal and informal savers with non-savers. We use our entire sample for this estimation. The estimate for Save formal (informal) shows the difference between the impact of Save formal (informal) and not saving (base category) on reinvestment likelihood. (2) In column 2, we compare the reinvestment performance of informal savers with formal saver by restricting our sample to only savers and dropping the respondents who do not save. The estimate for Save informal shows the difference between the impact of Save informal and Save formal (base category) on reinvestment likelihood. (3) In column 3, we disentangle informal saving practices to saving informal individually and saving informal with others by adding separate dummies for each group. The estimate for Save informal individually (Save informal with others) shows the difference between the impact of Save informal individually (Save informal with others) and Save formal (base category) on reinvestment likelihood. (4) In column 4, we compare reinvestment likelihood of household savers with formal savers by keeping only formal savers and household member savers in our sample. The estimate for Save informal with household members shows the difference between the impact of Save informal with household members and Save formal (base category) on reinvestment likelihood. (5) In column 5, we compare reinvestment likelihood of outside household savers with formal savers by keeping only formal savers and outside household savers in our sample. The estimate for Save informal with people outside household shows the difference between the impact of Save informal with people outside household and Save formal (base category) on reinvestment likelihood. (6) In column 6, we estimate the model shown in column 4 by using nearest neighbor matching estimator. 3 closest neighbors are used in the estimations. The biases regarding the continuous variables are corrected by bias adjustment procedure. p<0.1. ** p<0.05. *** p<0.01

Table 5: Test for the stability of estimates for Save with household members

	(1)	(2)	(3)	(4)	(5)	(6)
Save with household members	-0.15***	-0.14***	-0.12***	-0.13***	-0.13***	-0.12***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Borrowed		0.02	0.02	0.02	0.03	0.05
		(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Education			-0.01	-0.01	-0.01	-0.01
			(0.01)	(0.01)	(0.01)	(0.01)
Female				-0.08**	-0.09***	-0.07**
				(0.03)	(0.03)	(0.03)
Single				0.04	0.04	0.06
				(0.04)	(0.04)	(0.04)
No training			-0.04	-0.04	-0.03	-0.07**
			(0.03)	(0.03)	(0.03)	(0.03)
Income			0.04***	0.04***	0.04***	0.03**
			(0.01)	(0.01)	(0.01)	(0.01)
Observations	877	877	877	877	877	877
Pseudo R-squared	0.03	0.03	0.04	0.05	0.06	0.22
Sector FE	No	No	No	No	Yes	Yes
Region FE	No	No	No	No	No	Yes

Notes: In this table, we test the stability of our estimates for Save with household members shown in column 4 of Table 3 by inserting key control variable desperately. Reinvestment is the dependent variable in the estimations. We estimate Probit models for all specifications in columns 1 to 5 and report marginal effects estimates at mean values for all variables and robust standard errors in parentheses. Robust standard errors in parentheses. The definitions of the variables are given in Table 1. The details of the estimations in the columns 1 to 5 are as follows: (1) In column 1, we estimate the model withut a control variable. (2) In column 2, we control for access to finance by using our borrowing proxy (Borrowed). (3) In column 3, we control for productivity proxies, education, training, and income in addition to borrowing (4) In column 4, we additionally insert Female and Single variable to control for gender and being married. (5) In column 5, we control for sectoral heterogeneity by using sector dummies in addition previous control variables. (6) In column 6, we finally insert region dummies to control for regional heterogeneity. p<0.1. ** p<0.05. *** p<0.01

Table 6: Estimates for reinvestment and save with household members relationship-additional control variables

	(1)	(2)
Save with household members	-0.07**	-0.18***
	(0.03)	(0.05)
Formal Loan	0.02	
	(0.05)	
Semi formal loan	-0.11	
	(0.09)	
Informal loan	-0.12	
	(0.07)	
Initial capital stock	0.02**	
	(0.01)	
Sales per worker	-0.02	
	(0.01)	
Rural	0.03	
	(0.03)	
Size	0.04	
	(0.03)	
Duration	0.02	
	(0.01)	
Observations	872	650
Entrepreneurial dummies	Yes	No
Activity Dummy	Yes	No
Region FE	Yes	No
District FE	No	Yes

Notes: This table test the robustness of the estimates for Save informal with household members shown in Column 4 of Table 3 to the inclusion of additional control variables. Reinvestment is the dependent variable in the estimations. We estimate Probit models for the in columns 1 and 2 and report marginal effects estimates at mean values for all variables and robust standard errors in parentheses. In addition to the variables shown in the Table, we also add Borrowed, Education, Female, Single, No training, Income, and region dummies to the estimated models as control variables. The detailed definitions of these variables are given in Table 1. The sample used for estimations include only formal savers and informal savers with household members. Save formal is the base category for Save with household member estimates in both estimations. The details of the estimations in the columns 1 and 2 are as follows: (1) In column 1, we test whether the estimate for Save informal with household members is robust with respect to the inclusion of a vector of additional variables in addition to the listed variables above and region dummies. These additional control variables are as follows. Formal loan, Semi-formal loan, and Informal loan are dummy variables equals to 1 if the respondent received credit from formal, semi-formal and informal resources. Initial capital stock is the logarithm of the value of capital in U.S dollars invested to establish the business. Sales per worker is the sales divided by number of permanent workers of the business the respondent owns in logarithms. Rural is a dummy variable equals to 1 if entrepreneur lives at a rural area. Size is the number of permanent employees working in the business of the entrepreneur in logarithms. Duration is the logarithm of the number of years that the entrepreneur is doing the business. Entrepreneurial dummies consist of seven binary variables created for the listed answers to the question why did you go into business?. The list comprises statements I was fired / lost/retrenched from a previous job; I couldn't find a job elsewhere; To support me / my family; To try out a business idea; I believe I can make more money working for myself than for someone else; I had nothing else to do/no other means of survival/no better option; parents /relatives were in business; I saw a good opportunity; I have always wanted my own business; I was encouraged by friends and relatives; and I needed to supplement my income as an answer. Activity dummies consist of 5 dummy variables indicating whether the business owned by the respondent buy and sell goods; buy, add value and sell goods; make and sell goods; provide services; or do other activities including agricultural ones. (2) In column 2, we test whether the estimate for Save informal with household members is robust with respect to replacing region dummies with district dummies. Since there is no reinvestment variation at some district the number of observations is lower for this estimation. * p<0.1. ** p<0.05. *** p<0.01

Table 7: The estimates for save with household members vs. formal saving choices

	(1)	(2)
Bank branch withing	-0.13***	
one hour walking distance	(0.04)	
Min. distance to		0.04***
ATM, bank branch, or MFI		(0.01)
Age	-0.04***	-0.04***
	(0.01)	(0.01)
Age^2	0.00***	0.00***
	(0.00)	(0.00)
Borrowed	-0.32***	-0.32***
	(0.03)	(0.04)
Education	-0.12***	-0.12***
	(0.02)	(0.03)
Female	-0.10**	-0.10**
	(0.04)	(0.04)
Single	-0.14**	-0.14*
	(0.06)	(0.07)
No training	0.02	0.01
	(0.04)	(0.04)
Income	-0.10***	-0.10***
	(0.02)	(0.02)
Observations	877	797

Notes: This table shows the estimates for the determinants of the choice Save informal with household members vs. Save formal. Save informal with household members is the dependent variable, and Save formal is the base category for Save with household member estimates. The definitions for the variables are given in Table 1. We report marginal effects estimates at mean values for all estimations from Probit estimations and robust standard errors are in parentheses. The sample used for estimations include only formal savers and informal savers with household members. We additionally control for region fixed effects by adding region dummies to the estimations. The details of the estimations in column 1 and 2 are as follows: (1) In column 1, we use Bank branch within one hour walking distance as our distance formal banking services measure. (2) In column 2, we use Minimum distance to ATM, bank branch, or MFI as our distance formal banking services measure. p<0.1. ** p<0.05. *** p<0.01

Table 8: Tests for exogeneity of instruments and Bivariate Probit Estimates for save with household member

	Exogeneity checks		Bivariate Probit Estimates	
	(1)	(2)	(3)	(4)
Save with household members	-0.11***	-0.10***	-0.20**	-0.20**
	(0.04)	(0.04)	(0.10)	(0.10)
Bank branch within one hour	0.00			
walking distance	(0.03)			
Min. distance to ATM, bank		-0.01		
branch, or MFI		(0.01)		
Age	0.01	0.01		
	(0.01)	(0.01)		
Age^2	0.00	0.00		
	(0.00)	(0.00)		
$\hat{\mu}$			0.23	0.27
			(0.23)	(0.24)
Observations	877	797	877	797
Exogenous distance measure	-	-	Bank branch within one hour walking distance	Min. distance to ATM, bank branch, or MFI
Methodology	Probit	Probit	Bivariate Probit	Bivariate Probit

*Notes:*This table shows the bivariate probit estimates for Save informal with household members and summarize our informal exogeneity test for the excluded instruments. Reinvestment is the dependent variable for the estimates, and Save formal is the base category for Save with household member estimates. The sample used for estimations include only formal savers and informal savers with household members. We report marginal effect estimates of Save with household members at mean values, robust standard errors in the parentheses at columns 1 and 3, and clustered robust standard errors at ward level in the parentheses in columns 2-4. We also control for Borrowed, Education, Female, Single, No training, Income, and region dummies in all models. The detailed variable definitions are given in Table 1. The details of the estimations shown in Column 1 to 4 are as follows: (1) In column 1, we informally test whether our instruments, Bank branch within one hour walking distance, *Age*, and Age^2 do not have a direct impact on reinvestment likelihood in order to test the exogeneity of the instruments that will be used in Column 3 of the Table. If the estimates are not statistically significant than it may imply that they do not have direct impact on reinvestment but have an impact only through saving practice choice. Marginal effect estimates at mean values from Probit estimations and robust standard errors in the parenthesis are reported. (2) In column 2, we informally test whether our instruments, Minimum distance to ATM, bank branch, or MFI, *Age*, and Age^2 do not have a direct impact on reinvestment likelihood in order to test the exogeneity of the instruments that will be used in Column 4 of the Table. If the estimates are not statistically significant than it may imply that they do not have direct impact on reinvestment but have an impact only through saving practice choice. Marginal effect estimates at mean values from Probit estimations and robust standard errors in the parenthesis are reported. (3) In column 3, we report the marginal effect estimates for Save informal with household members from Bivariate Probit estimation. Robust standard errors are in parenthesis. Our excluded instruments for saving practice choice between Save with household members and Save formal are Bank branch within one hour walking distance, *Age*, and Age^2 .μ is the correlation estimate for the error terms from model 15 and 16. (4) In column 4, we report the marginal effect estimates for Save informal with household members from Bivariate Probit estimation. Robust standard errors are in parenthesis. Our excluded instruments for saving practice choice between Save with household members and Save formal are Minimum distance to ATM, bank branch, or MFI, *Age*, and Age^2. μ is the correlation estimate for the error terms from model 15 and 16. p<0.1. ** p<0.05. *** p<0.01

Table 9: Heterogeneity in the effect of saving with household members on reinvestment

	Gender		Position in the household	
	Male	Female	Other (Child, spouse, sibling etc)	Head
Save with household members	-0.12**	-0.22***	-0.22***	-0.16***
	(0.05)	(0.08)	(0.08)	(0.06)
Observations	402	275	213	441

Notes: This table summarizes the heterogeneity in the estimates for Save informal with household members for the benchmark estimate shown in column 4 of Table 3. To test the heterogeneity, we estimate the models in specific subsamples determined acccording to the respondent characteristics (please see below). Reinvestment is the dependent variable in the estimations, and Save formal is the base category for Save informal with household member estimates. The sample used for estimations include only formal savers and informal savers with household members. We report marginal effect estimates of the estimates at mean values from Probit estimations and robust standard errors in the parentheses. We also control for Borrowed, Education, Female, Single, No training, Income, sector and region dummies in all models. The detailed variable definition are given in Table 1. Since there is no reinvestment variation at some subsamples, the number of observations are lower for this estimation, and we therefore estimate all models for the same subsamples where there is variation in our reinvestment variables. The details of the estimations shown in Gender and Position in the household columns are as follows: (1) In Gender column, we estimate the model for the male and female respondents separately. (2) In Position in the household column, we separately estimate the model for the respondents who are household heads and others (not households such as child, spouse, parent, etc. of the household head). p<0.1. ** p<0.05. *** p<0.01

Appendix

Table A1: Estimates for reinvestment and saving/saving practices relationship by using full sample

VARIABLES	(1) reinvest	(2) reinvest	(3) reinvest
Save formal	0.09***	0.09***	0.09***
	(0.02)	(0.02)	(0.02)
Save informal	0.06***		
	(0.01)		
Save informal individually		0.07***	0.07***
		(0.01)	(0.01)
Save informal with others		0.01	
		(0.02)	
Save with household members			-0.01
			(0.02)
Save with people outside household			0.04
			(0.03)
Borrowed	0.04**	0.04**	0.04**
	(0.02)	(0.02)	(0.02)
Education	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)
Female	-0.03**	-0.03**	-0.03**
	(0.01)	(0.01)	(0.01)
Single	0.04**	0.04**	0.04**
	(0.02)	(0.02)	(0.02)
No training	-0.03***	-0.03***	-0.03***
	(0.01)	(0.01)	(0.01)
Income	0.03***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)
Observations	5,803	5,803	5,803
Sample	All	All	All
Hypothesis	p-values	p-values	p-values
H_0: Save formal-Save informal=0	0.0723	-	-
H_0: Save formal-Save informal individually=0	-	0.1819	0.1763
H_0: Save formal-Save informal with others=0	-	0.0011	-
H_0: Save formal-Save with household members=0	-	-	0.0004
H_0: Save formal-Save with people outside household=0	-	-	0.0662

Notes: This table shows the baseline estimation results for the relationship between saving practices, control variables and reinvestment likelihood by using the full sample for each estimation presented in columns 1 to 3. The detailed variable definitions are given in Table 1. Reinvestment is the dependent variable in the estimations. We estimate Probit models for all specifications and report marginal effects estimates at mean values for all variables and robust standard errors in parentheses. To control for unobserved regional and sector level fixed effects, we add sector and and region dummies to all estimations. The details of the estimations in the columns 1 to 3 are as follows: (1) In column 1, we compare the reinvestment likelihood of formal and informal savers with non-savers. The estimate for Save formal (informal) shows the difference between the impact of Save formal (informal) and not saving (base category) on reinvestment likelihood. (2) In column 2, we disentangle informal saving practices to saving informal individually and saving informal with others by adding separate dummies for each group. The estimate for Save informal individually (Save informal with others) shows the difference between the impact of Save informal individually (Save informal with others) and Not saving (base category) on reinvestment likelihood. (3) In column 3, we disentangle save informal with other to save informal with household members and save informal with people outside. The estimate for Save informal with household member (people outside) shows the difference between the impact of Save informal with household members (people outside) and Not saving (base category) on reinvestment likelihood. (4) We test the difference of the impacts of difference saving practices from Save formal at the bottom of the Table. H_0 indicates the null hypothesis of the test. p-values for the t-test are given in corresponding rows and columns. * p<0.1. ** p<0.05. *** p<0.01

Table A2: Bivariate Probit Estimates

Dependendent variable:	(1) Reinvestment	(2) Save with household members	(3) Reinvestment	(4) Save with household members
Save with household members	-0.81**		-0.86**	
	(0.37)		(0.41)	
Bank branch withing		-0.39***		
one hour walking distance		(0.12)		
Min. distance to				0.10***
Atm, bank branch, or MFI				(0.04)
Age		-0.12***		-0.11***
		(0.03)		(0.04)
Age^2		0.00***		0.00***
		(0.00)		(0.00)
Borrowed	0.07	-0.98***	0.05	-0.97***
	(0.16)	(0.12)	(0.17)	(0.13)
Education	-0.07	-0.33***	-0.07	-0.34***
	(0.06)	(0.07)	(0.07)	(0.07)
Female	-0.24**	-0.27**	-0.31***	-0.28**
	(0.12)	(0.12)	(0.12)	(0.13)
Single	0.25	-0.47**	0.25	-0.44*
	(0.18)	(0.22)	(0.20)	(0.24)
No training	-0.32***	0.05	-0.36***	0.04
	(0.12)	(0.12)	(0.14)	(0.12)
Income	0.08	-0.28***	0.06	-0.27***
	(0.06)	(0.05)	(0.06)	(0.05)
Constant	0.05	8.51***	0.56	8.08***
	(1.07)	(0.97)	(1.23)	(1.08)
Observations	877	877	797	797

Notes: This table shows the detailed bivariate probit coefficient estimates for the columns 3 and 4 of Table 5. We report bivariate probit estimates for all estimations. Robust standard errors for columns 1 and 3 and clustered robust standard errors at ward level in columns 2-4 are in parentheses. We use the sample for Formal Savers and Household Savers in all estimations, and Save formal is the base category for Save informal with household members. We additionally control for region dummies in the estimations. The details of columns 1 to 4 are as follows: (1) In column 1, we present the bivariate probit estimates of model (15) which is jointly estimated with model (16) using Bank branch within one hour walking distance as the distance measure. (2) In column 2, we present the bivariate probit estimates of model (16) using Bank branch within one hour walking distance as the distance measure and jointly estimated with model (15). (3) In column 3, we present the bivariate probit estimates of model (15) which is jointly estimated with model (16) using Minimum distance to ATM, bank branch, or MFI as the distance measure. (4) In column 2, we present the bivariate probit estimates of model (16) using using Minimum distance to ATM, bank branch, or MFI as the distance measure and jointly estimated with model (15). p<0.1. ** p<0.05. *** p<0.01